A Quantitative Approach for Web Usability Using Eye Tracking Data

López-Orozco and Florencia-Juárez

Additional information is available at the end of the chapter

Abstract

This chapter presents a relatively new approach to show how a web usability classical paradigm can benefit from quantitative data of a nonclassical approach. In the pilot stage, we used experimental eye tracking data acquired from 11 participants faced to a web page to perform three simple tasks. Results show advantages by using eye tracking data to identify and verify some usability problems of such a web page. In this chapter, some hints are presented for people interested in measuring web usability by using such an approach. However, a deeper study should be carried out in order to generalize our results toward the construction of a methodology to be followed by a web developer or interested people in such a field of research.

Keywords: eye tracking, web usability, human-computer interaction, quantitative approach, nonclassical approach

1. Introduction

Nowadays, usability is one of the key factors in the success or failure of a web project. Usability is considered a quality attribute that evaluates the accessibility, readability, navigability and ease of learning of a website or application by the user [1]. Moreover, people do not wish to learn how to use a website, and there is no manual to use a web page when a person does not find the information or products they are looking for within a website. Users should have the ability to understand how a website works immediately after viewing the website. If someone gets lost while browsing a website or has difficulty in reading the information, he or she simply tries not to use that site anymore [2].

Typically, usability has been measured in a qualitative way, using questionnaires, focus groups, interviews, and so on, and, to a lesser extent, in a quantitative manner by measuring time to complete a task, number of errors, number of clicks, and so on [3]. The qualitative methods offer the advantage of having a direct opinion of the users, such as what they like and dislike about the page and how it could be improved, but the opinions of the users could be different from each other. In contrast, the quantitative techniques help us to assign numerical values to statements or observations, in order to study the possible relationships between variables by using statistical methods.

It is important to emphasize that we do not consider any of the techniques mentioned earlier as one "best technique." None of them solve the problem of measuring usability completely. Both approaches have their advantages and disadvantages to consider a limited number of influencing factors [4]. However, techniques that combine the strengths of both are still better and more promising as shown here.

In this chapter, the eye tracking technique was used to obtain quantitative metrics based on what users see on a web page, with the aim of adding these metrics to the traditional qualitative observations, used in the measurement of usability of web pages.

2. Eye tracking technology

The eye tracker is a tool that detects the users' gaze and provides an accurate representation and understanding of an individual's eye movement behavior. An eye tracker detects the two basic eye movements: saccades and fixations. A saccade is the eye movement itself which is needed to place the sight in a different place and a fixation is the period of time when the eye is practically maintained on a single location.

Using an eye tracker allows:

- To see what the user observes on the screen by detecting the user's fixations, both the left and the right eyes on the X and Y coordinates according to the size of the screen as well as duration of each fixation.

- To know the route or scanpath that the user looks at. The eye tracker also provides the saccade route, which shows the fixations made and the movement between them as shown in **Figure 1**.

- To know the diameter of the pupil when making the fixations or at the moment, the user clicks the mouse on a certain position of the web page.

In this chapter, a binocular infrared S2 Eye Tracker Mirametrix Inc. was used. The eye tracker works with an accuracy of 0.5°, data rate of 60 Hz, a head movement box of 25 cm × 11 cm × 30 cm (width × height × depth); with a system setup time less than 5 min, calibration time less than 30 s; and a suggested screen size of 15"–22" monitor, with capacity of blink tracking recovery, fast time to track recovery, weight of 1 lb, eye tracking technique of bright pupil; and low-power computers.
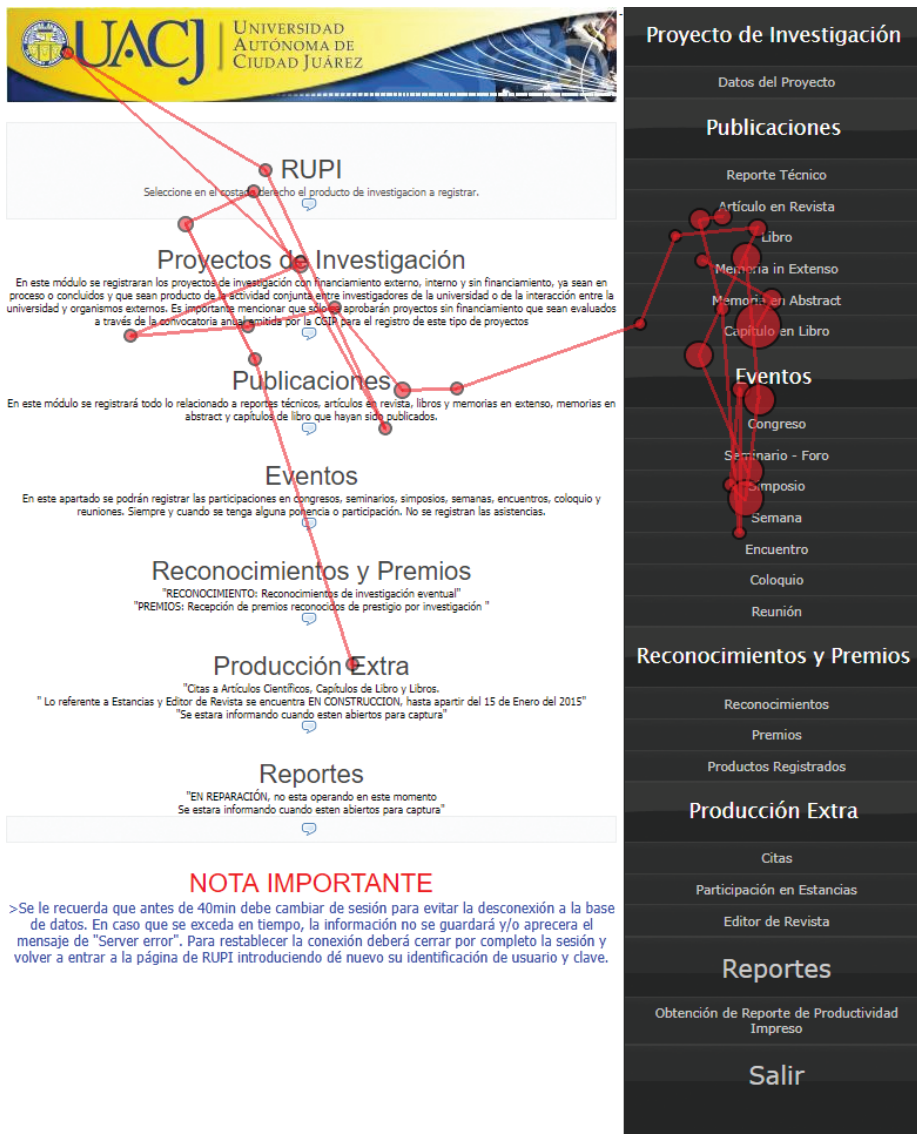
Figure 1. Saccade route generated by the eye tracker.

3. Related work

Some of the most representative research works that have used eye tracking data to solve usability problems in web pages are presented here.

The Centre for Human Computer Interaction Design of the City University of London [5] acquired experimental data by using a Tobii ×50 eye tracker. An ocular follow-up was performed while participants were doing tasks within two web pages. They found sequences of patterns of ocular tracking correlated with the sequences of patterns of usability problems, which are already known.

Similarly, a website for the American Society of Clinical Oncology (ASCO) was proposed, designed and evaluated with an eye tracking methodology [6]. Participants were asked to complete certain tasks, both in an old design of the page and in the new design. The paper describes how eye tracking helped to diagnose errors and identify the best of both designs.

Djamasbi et al. [7] conducted a study where they examined possible gender differences in web preference with the help of an eye tracker. Several studies show that both men and women have different esthetic preferences, for example, men prefer using darker colors, such as black and blue, while women prefer lighter colors. These preferences not only apply in colors but also in the website distribution of images and information. According to the previous studies, two hypotheses were presented: H1—participating women will notice more sections that have images of people more than men and H2—female participants will notice sections with a light-colored background more than men. The results did not show any significant difference between genders with respect to the number of times they were fixed in the sections of the web page. The results of the fixation analysis did not show significant differences between men and women with respect to sections with different background colors. Although the results did not support the hypotheses, the analyses yielded other data on the web page banners, suggesting that men and women can also ignore banner-type objects, regardless of whether these objects are designed considering their preferences.

Szymanski et al. [8] presented a method for evaluating web usability by combining kinetic gestures and eye tracking applications. All this was done by performing three tasks in an application that aims to deliver a list to users of all stores, restaurants and other places within a mall. The tasks performed were to find certain stores in a specific point in the commercial mall.

In the paper presented by Russell et al., at the annual British Human Computer Interaction Conference [9], it shows a study to discover the answer to two design issues both in online documents and on web pages. Research questions were: what is the best font size to read online? and what kind of source font is best for reading online? *Serif* or *Sans Serif*? The experiment was carried out with employees of a computer company, using an eye tracking device. Participants were given instructions to read and retain several texts with different sizes and types of source fonts, followed by answering three multiple-choice questions per text. Results show that larger fonts have a slight but not significant advantage in reading speed over small fonts. Results on the type of source between *Serif* and *Sans Serif* do not show significant differences in ocular follow-up and retention of texts.

Understanding the age-related differences in website navigation is instructive for website design, especially given by the growing number of older people using the Internet. Romano Bergstrom et al. [10] present the usability and data tracking of five independent usability studies of a website that included young and old participants. The results revealed age-dependent

differences in movement and eye functioning during website navigation in some of the sections. In particular, in one study, older participants were less precise, and in two studies, they took longer time to complete tasks compared to younger participants. In two studies, they observed the central part of the screen more frequently than younger participants and also observed the peripheral left side of the screen less frequently. Also, they took more time to cast first look at the peripheral top of the screen than the younger participants in the study. These results highlight the potential of age-related differences in performance while browsing websites and provide motivation for further exploration.

Tag-clouds are a compact list of keywords used in the web page interface, which allow users to browse and navigate through various documents or sections of those pages. In the article written by Montero et al. [11], a study was conducted to investigate the usability of tag-clouds. The experiment was performed in 17 participants, 6 women and 11 men aged between 29 and 49 years. Tasks asked to participants were divided into stages. First was a semantic search, where subjects were asked to search and click on a concept related to a certain topic, such as *Christianity*, where the expected target was the tag *religion*. The second stage consisted of a syntactic search, in which participants were asked to search and click on a tag that answered a certain question that was asked, such as searching for the name of the best-known operating system in the market, where the expected answer was *windows*. The results showed that both the size and the shape of the tag have a clear influence on the visual exploration of the users. They also showed results on the alphabetical order of the tags, which do not imply an improvement in terms of efficiency at the time of the location of the tags.

Most of the works reported here try to solve completely the usability problems either by using only eye tracking data or by a qualitative approach. In general, they only solve some aspects. We are convinced that by combining the best of both approaches (qualitative and quantitative) in usability measurements, results should be better. For this reason, this chapter proposes a combined approach of both.

4. Experimental approach

In order to study and get some data, we designed an experiment to face people to a simple task on a web page. The web page **R**egistro **Ú**nico de **P**roductos de **I**nvestigacion (RUPI) of the Universidad Autónoma de Ciudad Juárez (UACJ) was used as it is shown in **Figure 2**. During the experiment, participants were asked to fill forms with informational data provided to them beforehand. An eye tracker was used to get data from the user's gaze while conducting the experiment.

4.1. Participants

A total of nine full-time and two half-time professors of the Universidad Autonóma de Ciudad Juarez participated in the experiment. They were four women and seven men, aged between 28 and 43 years, and all of them are from the Electrical Engineering and Computer Science

Figure 2. RUPI home page.

Department. All participants had normal or corrected-to-normal visual acuity and had no known neurological disorders. They were naive with respect to the purpose of the study, and they gave written consent before the experiment.

4.2. Apparatus

Eye movements were recorded by using a remote binocular infrared S2 Eye Tracker, Mirametrix Inc., as mentioned in Section 2. The eye tracking system was used in the pupil corneal reflection tracking mode sampling at 60 Hz and 0.5–1 of accuracy. We ran our experiment by using the Viewer Software provided by the eye tracker manufacturer. Due to the fact that the eye tracker generates a special record of eye position and synchronization events, we had to perform a preprocessing stage in order to adequate this output to a more accessible format for our further analysis. A recording of both eyes' movements was done during the experiment; however, we are only concerned with the analysis of movement data of a single eye.

4.3. Training before the testing phase

Before starting the experiment, participants were trained on the tasks they would perform with an example of information to be managed during the testing phase. This step was done in an intuitive way without using the stimulus page of the real experiment. Additionally, some extra indications were given to the participants. Firstly, they were informed about the possibility to enter fictitious data but as similar as possible to real data only if they did not remember the precise information and secondly, the experiment could be interrupted at any time if they could not find the section where the information should be entered or whether they might feel lost within the website.

4.4. Calibration

The participants were asked to sit in front of the PC monitor at approximately 65 cm. Then, a screen with calibration points was presented as shown in **Figure 3**. The participants should look at each of these calibration points. Subsequently, the percentage of calibration error is displayed on the screen. It is noteworthy that the percentage average error was considered acceptable if it was less than 20%. **Figure 3** shows an example of calibration result considered not acceptable (average error: 35.1 > 20%) for our study purposes.

4.5. Experiment

After the training and calibration were completed, the participants were instructed to start the experimental session, in which they should complete the requested task. The initial screen was the home page of RUPI, where the participants had to find the section to enter the information that was given beforehand. The session ended when the participant pressed the button to save their data.

4.6. End of session

Once the session was finished, the participants were given a short survey about their experience with the website. This survey had questions concerning the graphical user interface and navigation on the page. They were also asked for any kind of help for future users of the website. Participants were also informed about the purpose of the experiment, and they had the opportunity to ask any question about our research project.
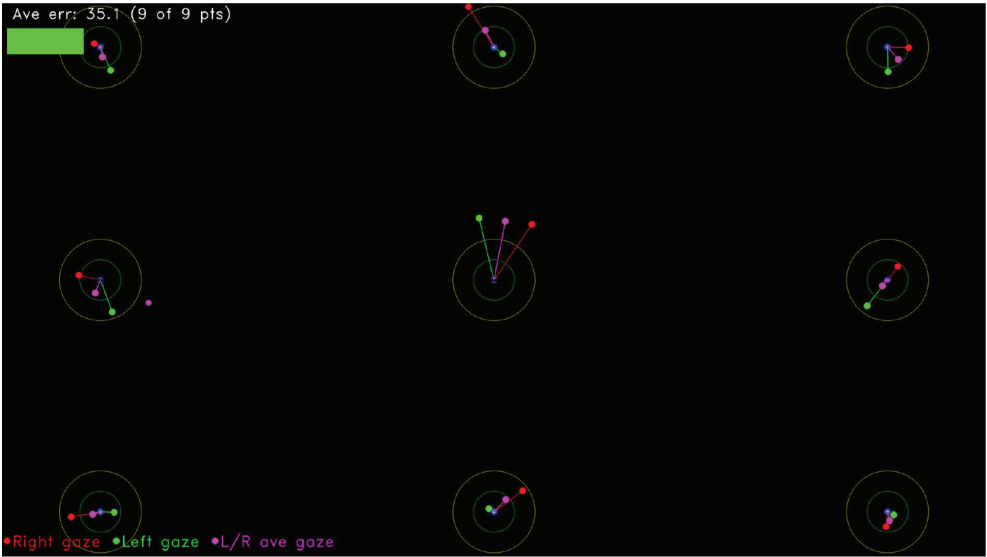
Figure 3. Calibration screen.
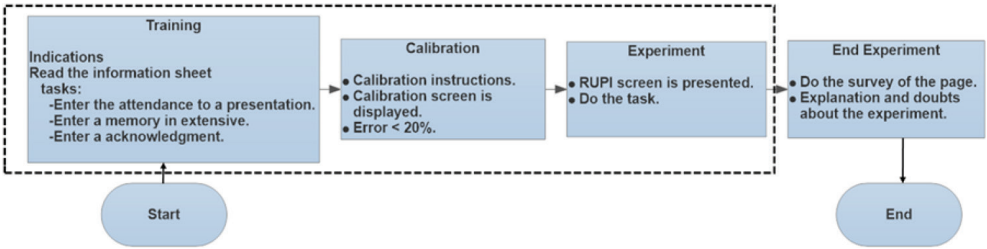
Figure 4. Overall phases performed during the experimental session. Phases bounded by dotted line were repeated three times by participants.

Figure 4 shows all the phases that were performed during the experimental session. Phases that are inside the dotted line are repeated three times for each task performed by the participant.

5. Data analysis

The main idea was to find usability problems by looking for regions or areas on the RUPI web page where participants made a lot of fixations. Our hypothesis was that if participants made a lot of fixation somewhere, this may be an indication of such a region which should be analyzed because people found it as being not clear or similar.

In this section, some statistics were computed from the experimental data. These computations are shown in **Table 1**.

P	Tj	TimeTj (s)	Fix	Order	OverallTime (s)	TotalFix	Fix/TimeTj (fix/s)	TotalFix/OverallTime (fix/s)
	1	154.95	9393	1			60.62	
P1	2	123.40	7487	2	358.93	21762	60.67	60.63
	3	80.58	4882	3			60.59	
	1	149.98	8862	2			59.09	
P2	2	161.40	9675	3	405.29	24242	59.94	59.81
	3	93.91	5705	1			60.75	
	1	125.81	7650	3			60.81	
P3	2	102.24	6212	1	364.03	21907	60.76	60.18
	3	135.98	8045	2			59.16	
	1	138.23	8398	2			60.75	
P4	2	199.90	11838	1	439.13	20236	59.22	46.08
	3	101.00	null	3			null	
	1	102.32	6213	2			60.72	
P5	2	528.84	31563	3	765.91	45966	59.68	60.01
	3	134.75	8190	1			60.78	
	1	94.48	5715	3			60.49	
P6	2	86.76	5192	2	292.33	17618	59.84	60.27
	3	111.09	6711	1			60.41	
	1	215.67	13079	1			60.64	
P7	2	285.65	17059	3	693.57	41827	59.72	60.31
	3	192.25	11689	2			60.80	
	1	118.92	7211	2			60.64	
P8	2	274.46	16635	1	594.09	36008	60.61	60.61
	3	200.71	12162	3			60.59	
	1	203.08	12340	1			60.76	
P9	2	370.39	20587	3	813.76	47172	55.58	57.97
	3	240.29	14245	2			59.28	
P10	1	282.24	17150	1	282.24	17150.00	60.76	60.76
P11	1	145.44	8847	1	145.44	8847.00	60.83	60.83

		First task performed by participant

Table 1. Computation of statics with experimental data.

5.1. Statistics

Eqs. (1)–(3) were used to compute the elapsed time and total number of fixations in each task performed by participants and a ratio between such number of fixations and the elapsed time, respectively.

$$OverallTime = \sum_{j=1}^{3} Time_{T_j} \qquad (1)$$

where $Time_{T_j}$ is the elapsed time (s) in each task performed by each participant,

$$TotalFix = \sum_{j=1}^{3} Fix_{T_j} \qquad (2)$$

where Fix_{T_j} is the number of fixations in each task performed by each participant and

$$\overline{Fix/Time} = \frac{TotalFix}{OverallTime} \tag{3}$$

giving the amount of fixations made for each participant per time unit. In other words, this variable is an indicator of "reading" speed.

5.1.1. Average time of the participants in completing the tasks

The average time to complete the first task for all participants was of 157.38 s ($\sigma = 55.77$), and for the second task, it was 237.00 s ($\sigma = 144.68$) and the average of the third task was 132.17 s ($\sigma = 72.50$).

It is important to mention that only two participants performed one task. For statistics computation, their data were aggregated to the other participants who did three tasks, so they were assigned as participant ID 10 and 11, respectively.

Additionally, **Figure 5** shows the number of fixations made in each task by participant, and the time of each participant in performing the three tasks was also computed and it is shown in **Figure 6**.

5.1.2. Scanpaths and selection of regions of interest

By using the geometric coordinates *(x, y)* and duration of the fixations, it is possible to draw the saccadic route of a participant in the web page. This saccadic route is known as scanpath. Once all the scanpaths of all participants were plotted, several regions of interest (ROIs) were selected by focusing for those with a high density of fixations.

Figure 7 shows the RUPI home page with the fixations of all scanpaths, and we only decided to analyze two ROIs in this interface: the *help menu* on the middle of the page (ROI$_1$) and the
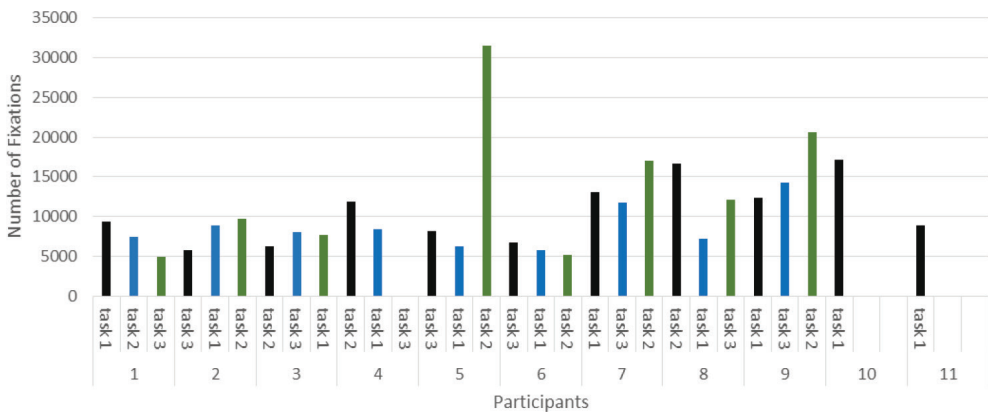

Figure 5. Number of fixations by participant. Tasks are presented in order performed.
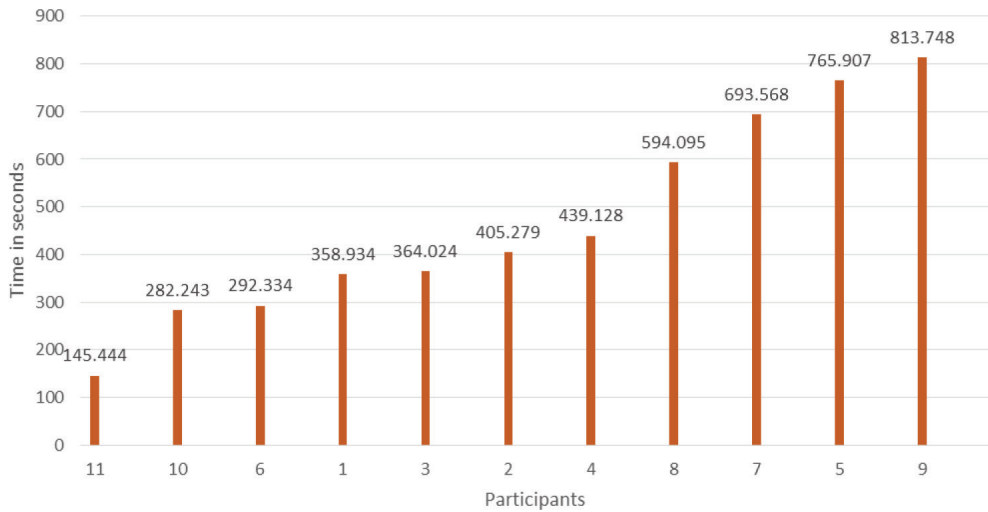
Figure 6. Time in completing the three tasks by participant ordered from lowest to highest.

right menu to enter the data on the right section (ROI$_2$), since those areas were the two with the highest fixation density as we can notice in such a figure.

Now, the qualitative data obtained from participants and how they could relate to the previous quantitative data are mentioned.

5.2. Qualitative data

After the experimental session, participants were asked to answer a 10-question short survey. This survey was intended in general to have information about the level of familiarization between the participant and RUPI web page, to know their experience during the experiment, to give their opinion about the graphical design of RUPI web page, and so on.

From the data obtained from the participant's survey (qualitative tests), the main results are described as follows:

- Most of the participants (8 of 11) have already once used the RUPI web page

- A total of 54.54% of the participants mentioned that the website's *help* section (ROI$_1$) to enter the data is not bad, and the other 45.4% mentioned that it is bad. This situation confirms our initial hypothesis that the website's *help* section is not clear, or it is not flashy

- A total of 45.45% of the participants had regular experience in finding the option to enter the information of the tasks, while 36.36% had bad experience. Only one participant could not enter the data and gave indications to end the session. It is important to mention that none of the participants entered the data in the corresponding section, since the RUPI web page is not well designed (ROI$_2$). For example, any information of a section can be entered in a different section.
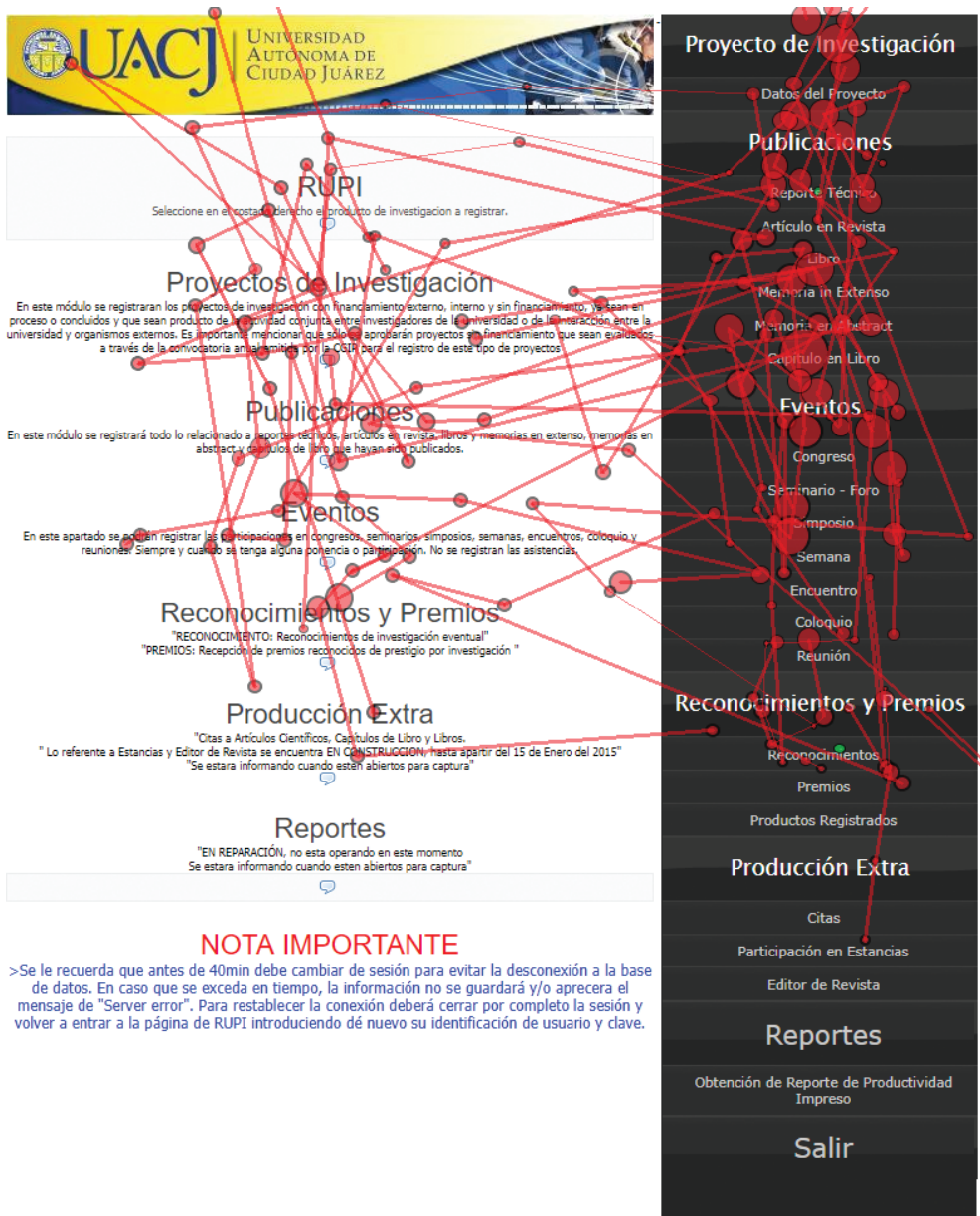
Figure 7. Scanpaths composed of fixations made by participants during the experiment.

5.3. Analysis of participants' behavior

In the ROI_2 located to the right of the web page, most of the participant fixations were observed when performing the very first task. Participants were concentrated in the *menu* area to enter

the options they were asked. Clearly, participants move a lot throughout the designated area as it is shown in **Figure 7**. This can give an indication that users get lost in the *menu* when trying to complete the first task.

Another point to mention is that there are very few fixations in the *help* area (ROI$_1$) of the website, giving way to understand several situations: (1) either this area was ignored often by users because it is not very flashy or (2) users who have already used the web page know that this section does not have clear information about the subsections and the data that should be entered in each one. This last assumption can be reinforced with the qualitative tests applied to the participants, since 54.54% of the participants mentioned that the website's *help* section to enter the data was not bad, and the other 45.45% mentioned that it was bad.

5.4. Redesigning and testing

Based on previous results and to test and confirm our findings quickly, simple modifications to the RUPI web page were proposed. Modifications to the visual and functional design were made. The scanpaths of participants testing our new design are shown in **Figure 8**. Here, fixations were better distributed among the different areas of the web page as it was expected.
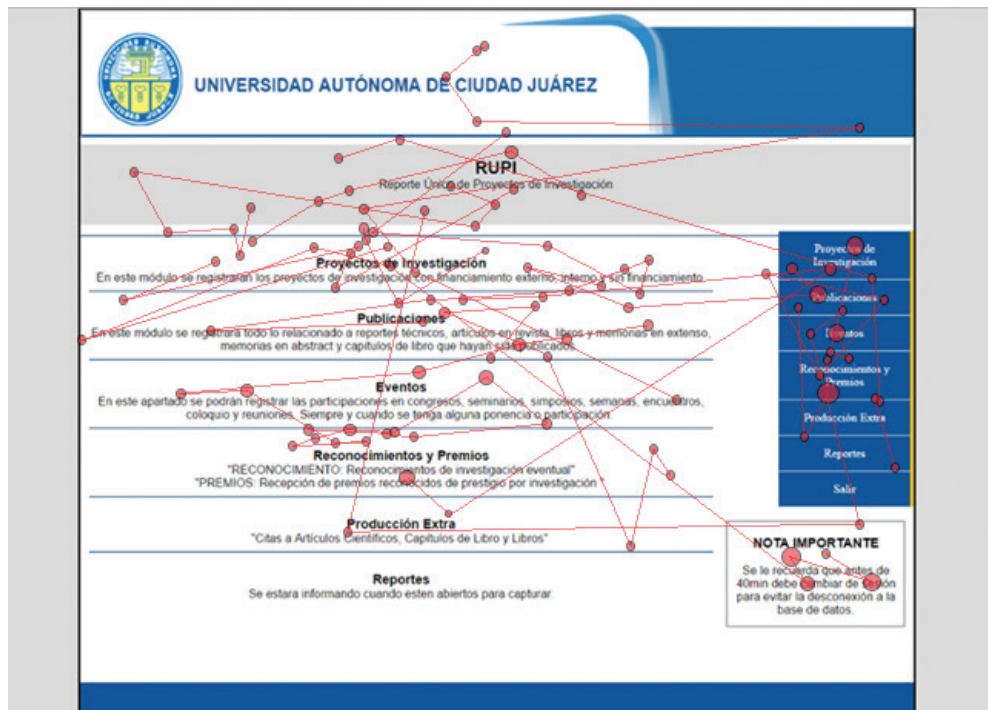

Figure 8. Scanpaths composed of fixations made by all participants during the experiment with the new web page design.

Moreover, 60% of participants rated the help section as good and 40% as regular. A total of 60% of participants entered the data successfully and 80% mentioned that they had not experienced any problem with the web page content. Finally, 60% of participants made plausible the fact that the new design has elements to allow users to go forward and go back while browsing.

6. Conclusions

In this chapter, a relatively new way of measuring usability by trying to combine the qualitative and quantitative techniques is proposed. Analysis of results showed how both techniques can be complemented. For example, an ROI where participants' rate was poor or marked as not helpful was *areas of help*. On the other hand, most of the works that try to solve usability problems by using the eye tracking technique generally do not focus on the qualitative part, as it could be seen in the state of the art, that is, qualitative measurement is not considered. More particularly, we have explored how the visual behavior of users can be combined with qualitative measurements to solve usability problems of the RUPI website. However, due to the complexity of human behavior and their reaction to a problem (in a web page) as well as the complexity of the required research, it is necessary to study this subject in more depth.

Acknowledgements

Thanks

Thanks to the student community involved during the experimental phase, specially to Mario Luna-Maldonado and Ramón G. Ortiz-Bustillos who carried out the experiment as part of their undergraduate project. Also, we thank the teaching staff of UACJ for accepting to participate in our experiments.

Author details

López-Orozco* and Florencia-Juárez

*Address all correspondence to: francisco.orozco@uacj.mx

Autonomous University of Ciudad Juárez, Ciudad Juárez, Mexico

References

[1] Sanchez W. La usabilidad en Ingeniería de Software: definicion y caracteríísticas, Ing-novacion. San Salvador: Universidad Don Bosco; Aug 2011

[2] Nielsen J. Designing web usability: The practice of simplicity. Thousand Oaks, CA, USA: New Riders Publishing; 1999

[3] Nielsen J, Pernice K. Eyetracking web usability. 1st ed. Thousand Oaks, CA, USA: New Riders Publishing; 2010

[4] Conte T, Vaz V, Massolar J, Mendes E, Travassos H. Improving a web usability inspection technique using qualitative and quantitative data from an observational study. In: Software Engineering, 2009. SBES'09. XXIII Brazilian Symposium. Fortaleza, Ceara, Brazil: IEEE; Oct 2009

[5] Ehmke C, Wilson S. Identifying web usability problems from eye-tracking data. In: Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI…But Not as We Know It. Vol. 1. British Computer Society. United Kingdom: University of Lancaster; Sep 2007. pp. 119-128

[6] Bojko A. Using eyetracking to compare web page designs: A case study. Journal of Usability Studies. 2006:**1**(3):112-120

[7] Djamasbi S, Tullis T, Hsu J, Mazuera E, Osberg K, Bosch J. Gender preferences in web design: Usability testing through eye tracking. In: Proceedings of the Thirteenth Americas Conference on Information Systems (AMCIS), Keystone, Colorado; August 2007. pp. 1-8

[8] Szymanski JM, Sobecki J, Chynał P, Anisiewicz J. Eye tracking in gesture based user interfaces usability testing. In: Asian Conference on Intelligent Information and Database Systems. Springer, Cham: Springer International Publishing; Mar 2015. pp. 359-366

[9] Beymer D, Russell D, Orton P. An eye tracking study of how font size and type influence online reading. In: Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction. Vol. 2. British Computer Society; Liverpool, United Kingdom. Sep 2008. pp. 15-18

[10] Romano Bergstrom JC, Olmsted-Hawala EL, Jans ME. Age-related differences in eye tracking and usability performance: Website usability for older adults. International Journal of Human-Computer Interaction. 2013;**29**(8)

[11] Montero YH, Herrero-Solana V, Guerrero-Bote V. Usabilidad de los tag-clouds: estudio mediante eye-tracking. In: Scire: Representacion y organizaci ´ on del conocimiento. Idioma: español; 2010;**16**(1):31-41. ISSN: 1135-3716